



Optimal Transport using Helmholtz-Hodge Decomposition and First-Order Primal-Dual Algorithms

Morgane Henry, Emmanuel Maitre, Valérie Perrier

► To cite this version:

Morgane Henry, Emmanuel Maitre, Valérie Perrier. Optimal Transport using Helmholtz-Hodge Decomposition and First-Order Primal-Dual Algorithms. 2015 IEEE International Conference on Image Processing (ICIP), Sep 2015, Quebec City, QC, Canada. pp.4748-4752, 10.1109/ICIP.2015.7351708 . hal-01134194

HAL Id: hal-01134194

<https://hal.science/hal-01134194>

Submitted on 23 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPTIMAL TRANSPORT USING HELMHOLTZ-HODGE DECOMPOSITION AND FIRST-ORDER PRIMAL-DUAL ALGORITHMS

Morgane Henry, Emmanuel Maitre and Valérie Perrier

University of Grenoble-Alpes
Laboratoire Jean Kuntzmann
St Martin d'Hères, France

ABSTRACT

This work deals with the resolution of the optimal transport problem between 2D images in the fluid mechanics framework of Benamou and Brenier formulation [1], which numerical resolution is still challenging even for medium-sized images. We develop a method using the Helmholtz-Hodge decomposition [2] in order to enforce the divergence-free constraint throughout the iterations. We then show how to use a first order primal-dual algorithm for convex problems of Chambolle and Pock [3] to solve the obtained problem, leading to a new algorithm easy to implement. Besides, numerical experiments demonstrate that this algorithm is faster than state of the art methods and efficient with real-sized images.

Index Terms— Convex optimization, optimal transport, proximal splitting, image processing, Helmholtz-Hodge decomposition

1. INTRODUCTION

Optimal transport is a domain with an increasing number of applications in economy [4], machine learning [5] or partial differential equations [6, 7]. The optimal transport problem defines a metric between densities [8], which appears to be relevant in image processing [9, 10]. The development of efficient new algorithms for the calculus of the optimal transport between two densities is still a challenge, especially for real-sized images. In this paper we are interested in the Benamou and Brenier formulation [1], who placed the problem in a context of fluid mechanics by adding a time dimension. They developed an algorithm based on the minimization of a functional which preserves the mass, using an augmented Lagrangian. Existing algorithms [10, 1], require a projection onto the divergence-free constraint at each iteration of the algorithm. This corresponds to solve a 3D Poisson equation at each time step for a 2D image. To reduce the computational cost, we decided to work directly in the space of constraints for the functional to minimize. Indeed, this will get

rid of solving the Poisson equation, and speeding up the algorithm. The preservation of the constraint will be given by the Helmholtz-Hodge decomposition [2] of divergence-free velocities, applied to the vector formed by the time-dependent densities and momentum. This allows to rewrite the functional of Benamou-Brenier in terms of a stream function, and we show that the first order primal-dual algorithm for convex problems of Chambolle Pock [3] can be easily adapted for finding the minimum of the new functional. The Chambolle-Pock method is nowadays widely used [11, 12], leading to fast implementations, since it can be easily speed up on parallel architectures. Therefore our method leads to a fast algorithm, simple to implement on imaging problems.

In the following, we begin by introducing the optimal transport framework in the first section. Then, we develop the decomposition we use to stay in the set of constraints. Afterward, we apply a primal-dual algorithm dedicated to our functional. We finish by numerical experiments, comparing our algorithm to state of the art on several test cases.

2. THE L^2 MONGE-KANTOROVICH PROBLEM

Let $\Omega = (0, 1)^2$ and $(\rho_0, \rho_1) \in L^2(\Omega)$, be two positive, bounded densities with $\int_{\Omega} \rho_0 = \int_{\Omega} \rho_1 = 1$. Let $|\cdot|$ be the Euclidean norm in \mathbb{R}^2 , the L^2 -Wasserstein distance (see for example [8]) between ρ_0 and ρ_1 is defined by

$$d_2(\rho_0, \rho_1)^2 = \inf_M \int |M(x) - x|^2 \rho_0(x) dx,$$

where the infimum is taken among the maps M transferring ρ_0 to ρ_1 , which means that $\forall A \subset \Omega, \int_{x \in A} \rho_1 = \int_{M(x) \in A} \rho_0$. The Monge-Kantorovich problem (MKP) amounts to determine an application M which realizes this infimum. Benamou and Brenier [1] rephrased the problem in a continuum mechanics framework. Let consider a time interval $(0, 1)$, we set $Q = (0, 1) \times \Omega$ and $V(Q) = \{f \in (L^2(Q))^{1+2}, \operatorname{div}_{t,x} f = 0\}$. We consider the densities $\rho(t, x) \geq 0$ and vector fields $v(t, x) \in \mathbb{R}^2$ verifying the continuity equation

$$\operatorname{div}_{t,x}(\rho, \rho v) = \partial_t \rho + \operatorname{div}_x(\rho v) = 0 \quad (1)$$

This work is supported by the French Agence Nationale de la Recherche (ANR, Project TOMMI) under reference ANR-11-BS01-014-01.

for $t \in (0, 1)$ and $x \in \Omega$, equipped with the initial, final and boundary conditions

$$\begin{aligned} \rho(0, x) &= \rho_0(x), \quad \rho(1, x) = \rho_1(x), \quad \forall x \in \Omega, \\ \rho v(t, x) \cdot \nu_\Omega &= 0, \quad \forall t \in (0, 1), \quad x \in \partial\Omega, \end{aligned} \quad (2)$$

where ν_Ω is the outward normal of Ω . As proven in [1] (see also [8]), the square of the L^2 -Wassertein distance between ρ_0 and ρ_1 verifies

$$d_2(\rho_0, \rho_1)^2 = \inf \int_0^1 \int_\Omega \rho(t, x) |v(t, x)|^2 dx dt,$$

where the infimum is taken among all ρ, v satisfying (1) and (2). To obtain a convex problem with linear constraints, Benamou and Brenier introduced the momentum $m = \rho v$ and obtained the following formulation

$$\min_{(\rho, m) \in C} \mathcal{J}(\rho, m) = \min_{(\rho, m) \in C} \int_0^1 \int_\Omega J(\rho(t, x), m(t, x)) dx dt,$$

where

$$\forall (\rho, m) \in \mathbb{R} \times \mathbb{R}^2, \quad J(\rho, m) = \begin{cases} \frac{|m|^2}{2\rho}, & \text{if } \rho > 0, \\ 0, & \text{if } (\rho, m) = (0, 0), \\ +\infty, & \text{otherwise,} \end{cases} \quad (3)$$

and the affine space of constraints reads

$$C := \{(\rho, m), \operatorname{div}_{t,x}(\rho, m) = 0, m(\cdot, x) \cdot \nu_\Omega = 0, \forall x \in \partial\Omega, \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1\}.$$

We present below an algorithm working directly in the set of constraints C .

3. REFORMULATION OF THE PROBLEM USING HELMHOLTZ-HODGE DECOMPOSITION

To work in C , we use the orthogonal decomposition of $L^2(Q)^{1+2}$, detailed in [2]. The vector field $v = (\rho, m) \in V(Q)$ has the following Helmholtz-Hodge decomposition:

$$(\rho, m) = \nabla \times \phi + \nabla h,$$

where we will denote $\nabla = \nabla_{t,x}$, in the following. Moreover $\phi \in (H_0^1(Q))^3$, and $h \in H^1(Q)/\mathbb{R}$ and we have also $\operatorname{div} \phi = 0$. Because (ρ, m) is divergence-free we obtain

$$\begin{cases} \Delta h = 0 \text{ in } Q, \\ \frac{\partial h}{\partial \nu_Q} = (\rho, m) \cdot \nu_Q \text{ on } \partial Q, \end{cases} \quad (4)$$

where ν_Q is the outward normal of Q . So we have first to solve the system (4) to obtain h , which is no more than a Poisson equation with known boundary conditions. Then, knowing h , we have to find the minimum of our new energy

$$\mathcal{J}(\nabla \times \phi) = \int_0^1 \int_\Omega F(\nabla \times \phi(t, x) + \nabla h(t, x)) dx dt, \quad (5)$$

where $F : (X, Y) \mapsto \frac{|Y|^2}{2X}$.

4. FIRST ORDER PRIMAL-DUAL ALGORITHM

The method described by Chambolle and Pock in [3], allowing to minimize energies of the form (5), uses a primal-dual formulation (see [13]) of the form:

$$\min_{\phi} \max_z \langle K\phi, z \rangle + \iota_{C_0}(\phi) - \mathcal{J}^*(z). \quad (6)$$

We consider $K = \nabla \times$, the curl operator, which is a linear continuous operator from $(H^1(Q))^3$ to $(L^2(Q))^3$, the Legendre transform of \mathcal{J} (see [14]) $\mathcal{J}^* : (L^2(Q))^3 \rightarrow [0, +\infty)$ and $\iota_{C_0} : (H^1(Q))^3 \rightarrow [0, +\infty)$, the indicator function of the set $C_0 := \{(\rho, m), m(\cdot, x) \cdot \nu_\Omega = 0, \forall x \in \partial\Omega, \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1\}$, which are proper, convex, lower semicontinuous functions. It has been shown in [15] that for $\theta = 1$ and $\sigma\tau\|K\|^2 < 1$, ϕ^k computed with the following algorithm, converges to the solution of (6):

Algorithm 1.

Initialization: $\tau, \sigma > 0, \theta \in [0, 1]$, $(\phi^0, z^0 = K\phi^0, \tilde{\phi}^0 = \phi^0)$.

Iterations:

$$\begin{aligned} z^{k+1} &= \operatorname{prox}_{\mathcal{J}^*}(z^k + \sigma(K\tilde{\phi}^k)) \\ \phi^{k+1} &= \operatorname{prox}_{\iota_{C_0}}(\phi^k - \tau K^* z^{k+1}) \\ \tilde{\phi}^{k+1} &= \phi^{k+1} + \theta(\phi^{k+1} - \phi^k). \end{aligned}$$

Detailing the steps of the algorithm. The discrete objective functional \mathcal{J} reads for (ρ, m) defined on the centered grid G^c (defined in section 5):

$$\mathcal{J}(\rho, m) = \sum_{k \in G^c} J(\rho_k, m_k), \quad (7)$$

where the functional J is defined in (3), and then,

$$\operatorname{prox}_{\gamma \mathcal{J}}(x) = (\operatorname{prox}_{\gamma J}(x_k))_{k \in G^c}.$$

As proved in [1], the Legendre transform of \mathcal{J} is the indicator function of a convex set, $\mathcal{J}^* = i_{P_J}$ where

$$\begin{cases} P_J = \{(z_1, z_2); \forall k \in G^c, (z_1, z_2)_k \in P_J\} \\ P_J = \{(t, x) \in \mathbb{R} \times \mathbb{R}^2, t + \frac{|x|^2}{2} \leq 0\}. \end{cases}$$

This implies that $\operatorname{prox}_{\gamma \mathcal{J}^*}$ is the projection onto the paraboloid P_J , which we will denote by \mathcal{P}_{P_J} . As we work on the constraint set C , $(\rho, m) = \nabla \times \phi + \nabla h$, we now define a new functional, for $(a, b) = \nabla \times \phi$:

$$\begin{aligned} \mathcal{J}_h(a, b) &= \mathcal{J}(a + \partial_t h, b + \nabla_x h) \\ &= \sum_{k \in G^c} J_h(a, b) = \sum_{k \in G^c} J(a + \partial_t h, b + \nabla_x h). \end{aligned} \quad (8)$$

This enables us to deduce from J^* the form of J_h^* and the form of $\operatorname{prox}_{\gamma J_h^*}$ from the one of $\operatorname{prox}_{\gamma J^*}$. If we denote $c = (a, b)$ we have the following proposition:

Proposition 1. One has for all $c \in \mathbb{R}^{1+n}$

$$J_h^*(c) = J^*(c) - \langle \nabla h, c \rangle,$$

and

$$\text{prox}_{\gamma J_h^*}(c) = \text{prox}_{\gamma J^*}(c - \gamma \nabla h).$$

Finally, the primal-dual algorithm reads in our case

Algorithm 2.

Initialization: $\tau, \sigma > 0, \theta \in [0, 1], (\phi^0, z^0 = K\phi^0, \tilde{\phi}^0 = \phi^0)$.

Iterations:

$$\begin{aligned} z^{k+1} &= \mathcal{P}_{P_J}(z^k + \sigma(\nabla \times \tilde{\phi}^k + \nabla h)) \\ \phi^{k+1} &= \mathcal{P}_{C_0}(\phi^k - \tau \nabla^* \times z^{k+1}) \\ \tilde{\phi}^{k+1} &= \phi^{k+1} + \theta(\phi^{k+1} - \phi^k). \end{aligned}$$

The computation of \mathcal{P}_{P_J} amounts to solve a third order equation by grid point, while \mathcal{P}_{C_0} merely corresponds to set the boundary conditions to zero.

5. NUMERICAL APPLICATION TO IMAGE TRANSPORT

5.1. Discrete setting

We now describe the discrete grids used in the computations. **Centered grid.** The regular grid

$$G^c = \{t_i, x_j, y_k\}_{1 \leq i \leq M, 1 \leq j \leq N, 1 \leq k \leq P},$$

with $t_i = \frac{i}{M}$, $x_j = \frac{j}{N}$, $y_k = \frac{k}{P}$ the discrete locations in the domain Q , is used to evaluate ρ and m .

Staggered grid. We introduce two staggered grids to evaluate the divergence and the curl operators. The grid G^{s1} provides a discretization coherent with the divergence of (ρ, m) and is defined by:

$$\begin{aligned} G_t^{s1} &= \{t_{i-1/2}, x_j, y_k\}_{1 \leq i \leq M+1, 1 \leq j \leq N, 1 \leq k \leq P}, \\ G_x^{s1} &= \{t_i, x_{j-1/2}, y_k\}_{1 \leq i \leq M, 1 \leq j \leq N+1, 1 \leq k \leq P}, \\ G_y^{s1} &= \{t_i, x_j, y_{k-1/2}\}_{1 \leq i \leq M, 1 \leq j \leq N, 1 \leq k \leq P+1}. \end{aligned}$$

Our staggered grid G^{s2} for ϕ such that $\nabla \times \phi$ is on the staggered grid G^{s1} :

$$\begin{aligned} G_t^{s2} &= \{t_i, x_{j-1/2}, y_{k-1/2}\}_{1 \leq i \leq M, 1 \leq j \leq N+1, 1 \leq k \leq P+1}, \\ G_x^{s2} &= \{t_{i-1/2}, x_j, y_{k-1/2}\}_{1 \leq i \leq M+1, 1 \leq j \leq N, 1 \leq k \leq P+1}, \\ G_y^{s2} &= \{t_{i-1/2}, x_{j-1/2}, y_k\}_{1 \leq i \leq M+1, 1 \leq j \leq N+1, 1 \leq k \leq P}. \end{aligned}$$

Interpolation operator. To go to the centered grid from the grid G^{s1} we need an interpolation operator, which is:

$$\begin{aligned} \rho, m(G^{s1}) &\rightarrow \rho, m(G^c) \\ \rho_{i-1/2,j,k} &\rightarrow \rho_{i,j,k} = (\rho_{i+1/2,j,k} + \rho_{i-1/2,j,k})/2 \\ m_{i,j-1/2,k} &\rightarrow m_{i,j,k}^1 = (m_{i,j+1/2,k}^1 + m_{i,j-1/2,k}^1)/2 \\ m_{i,j,k-1/2} &\rightarrow m_{i,j,k}^2 = (m_{i,j,k+1/2}^2 + m_{i,j,k-1/2}^2)/2 \end{aligned}$$

and its adjoint operator to go from G^c to G^{s1} .

Curl, gradient and divergence operators. The discrete gradient is a vector of matrices $\nabla v = (\nabla_t v \nabla_{x_1} v \nabla_{x_2} v)$. We use finite differences to compute the gradient, which has, for first component

$$\nabla_t v_{i,j,k} = v_{i+1/2,j,k} - v_{i-1/2,j,k}, \quad \text{if } i \leq M,$$

and the adjoint divergence operator

$$\nabla_t \cdot v_{i-1/2,j,k} = \begin{cases} -v_{1,j,k} & \text{if } i = 1 \\ v_{i,j,k} - v_{i-1,j,k} & \text{if } 2 \leq i \leq M \\ v_{M,j,k} & \text{if } i = M + 1. \end{cases}$$

The curl operator we use is derived from the gradient operator. In order to use the primal dual algorithm we need to define the discrete adjoint operator of the curl. Because the curl operator is given by the following matrix

$$\begin{pmatrix} 0 & -\nabla_{x_2} & \nabla_{x_1} \\ \nabla_{x_2} & 0 & -\nabla_t \\ -\nabla_{x_1} & \nabla_t & 0 \end{pmatrix}$$

the appropriate adjoint curl operator has to be the opposite of the curl derived from the divergence operator.

5.2. Numerical applications

For the performance evaluation we compare our algorithm (PDHH) to the primal-dual algorithm developed in [10] that we will denote PDPOP in the following. We computed $it = 10^6$ iterations in the case of the transport of two isotropic Gaussians with the same variance, and we plot the estimated density in Figure 1: the solution is displayed in black and grey, black being 0 and white being 1, and will be denoted (ρ_s, m_s) . We use a grid of $N \times P = 64 \times 64$ discretization points for ρ_0 and ρ_1 and $M = 64$ points for the time t . In the following we will use the parameters $\|K\|^2 = 8, \sigma = 90, \tau = 0.99/L\sigma$ and $\theta = 1$. We choosed σ such that the errors on m and ρ are minimal after $it = 50$ iterations.

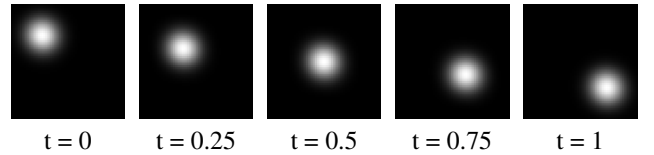


Fig. 1. Display of the density $\rho(t)$ obtained after $it = 10^6$ iterations.

Figure 2 displays the L^2 error between ρ and ρ_s and between m and m_s , the functional \mathcal{J} and the complexity, for 5000 iterations, for both algorithms PDPOP and PDHH in the example of Figure 1. It shows that despite our algorithm has not the best convergence rate during the first iterations, it converges quickly until we obtain the $O(1/i)$ convergence rate. Indeed,

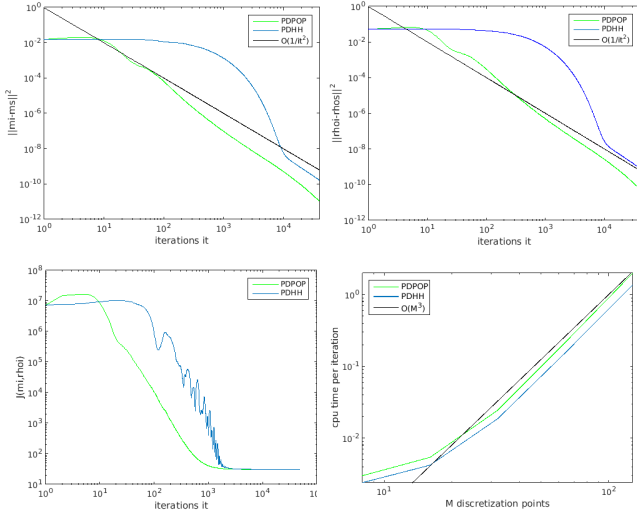


Fig. 2. Comparison at each iteration of the L^2 error between ρ and ρ_s and between m and m_s , the functional \mathcal{J} and the complexity, for 5000 iterations, for PDPOP algorithm and PDHH algorithm in the case of figure 1.

the decreasing of the functional in the constraint set has not the same behavior as in the PDPOP algorithm, where one has to project onto the divergence-free constraint space. Figure 2 also displays the computation time with respect to the number $M = N = P$ of discretization points in one direction. It shows that the complexity of the two algorithms is linear of order $O(M^3)$. But it depends also on the number of iterations. The bigger the grid is, the better our algorithm behaves in comparison with the PDPOP algorithm. Moreover, this behavior increases with the number of iterations we run, as shown in Table 1.

The explanation is that we don't have to solve a Poisson

	$it = 100$	$it = 500$	$it = 1000$	$it = 5000$
$N = 16$	1.26	1.24	1.20	1.26
$N = 32$	1.28	1.41	1.45	1.31
$N = 64$	1.15	1.31	1.37	1.40
$N = 128$	1.32	1.21	1.42	1.46

Table 1. Ratios between cpu time per iteration for PDPOP algorithm and PDHH algorithm, for different numbers of iterations and different sizes.

equation at each iteration. But contrarily to PDPOP, we have to evaluate a curl operator in K , which is slightly time-consuming.

Test on non convex densities. The next example of transport considers the case of irregular, non convex and non connected densities with compact support. Figure 3 shows the ability of our method to estimate the density $\rho(t)$ for such initial and final densities.

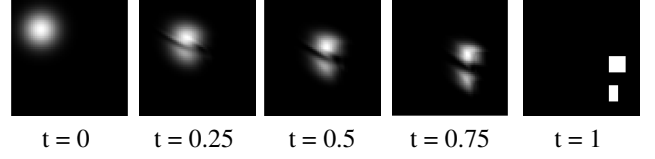


Fig. 3. Display of the density $\rho(t)$ obtained after $it = 10^6$ iterations of a non-convex, non connected density with compact support on a grid $M \times N \times P = 64 \times 64 \times 64$.

Test on real images. As last example we compute the density $\rho(t)$ for images representing clouds in different positions. The results presented in Figure 4 are obtained for images discretized on a grid $M = 30$ for the time dimension and $N \times P = 100 \times 68$ for the space dimension.

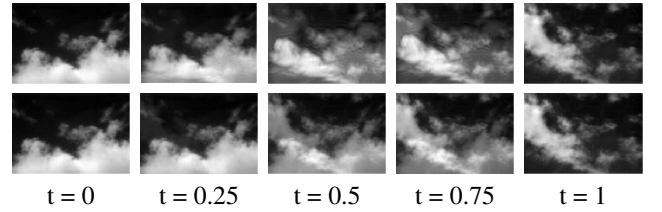


Fig. 4. Display of the density $\rho(t)$ of an image of clouds. The first line represents $\rho(t)$ obtained after $it = 1.10^6$ iterations of PDHH algorithm while the second line represents the L^2 interpolation.

6. CONCLUSION

We introduced a new algorithm for the optimal transport problem between 2D images, which respects the divergence-free constraint throughout the iterations, and therefore gets rid of solving a 3D Poisson equation at each iteration. Besides, this algorithm is easy to implement, faster than state of the art methods, and efficient for real-sized images. Further improvements of the method will include other divergence-free decomposition, and other formulations of the primal-dual algorithm.

7. REFERENCES

- [1] J-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the monge-kantorovich mass transfer problem," *Numerische Mathematik*, vol. 84, no. 3, pp. 375–393, 2000.
- [2] V. Girault and P-A. Raviart, "Finite element methods for navier-stokes equations: theory and algorithms," *NASA STI/Recon Technical Report A*, vol. 87, pp. 52227, 1986.
- [3] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imag-

- ing,” *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.
- [4] P-A. Chiappori, R. J. McCann, and L. P. Nesheim, “Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness,” *Economic Theory*, vol. 42, no. 2, pp. 317–354, 2010.
- [5] Y. Rubner, C. Tomasi, and L. J. Guibas, “A metric for distributions with applications to image databases,” pp. 59–66, 1998.
- [6] J. S. Moll J. A. Carrillo, “Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms,” *SIAM J. Sci. Comput.*, vol. 31, pp. 4305–4329, 2009.
- [7] R. Jordan, D. Kinderlehrer, and F. Otto, “The variational formulation of the fokker–planck equation,” *SIAM journal on mathematical analysis*, vol. 29, no. 1, pp. 1–17, 1998.
- [8] Villani C., *Topics in optimal transportation*, Number 58. American Mathematical Soc., 2003.
- [9] J. Rabin, S. Ferradans, and N. Papadakis, “Adaptive color transfer with relaxed optimal transport,” pp. 4852–4856, 2014.
- [10] N. Papadakis, G. Peyré, and E. Oudet, “Optimal transport with proximal splitting,” *SIAM Journal on Imaging Sciences*, vol. 7, no. 1, pp. 212–238, 2014.
- [11] J. C. Pesquet P. L. Combettes, L. Condat and B. C. Vu, “A forward-backward view of some primal-dual optimization methods in image recovery,” pp. 4141–4145, 2014.
- [12] B. He and X. Yuan, “Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective,” *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 119–149, 2012.
- [13] R T. Rockafellar, *Convex analysis*, vol. 28, Princeton university press, 1997.
- [14] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, 2011.
- [15] A. Chambolle, “An algorithm for total variation minimization and applications,” *Journal of Mathematical imaging and vision*, vol. 20, no. 1-2, pp. 89–97, 2004.
- [16] K. J. Arrow, L. Hurwicz, H. Uzawa, and H. B. Chenery, *Studies in linear and non-linear programming*, vol. 2, Stanford University Press Stanford, 1958.
- [17] P. L. Combettes and V. R. Wajs, “Signal recovery by proximal forward-backward splitting,” *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.
- [18] R. Yildizoglu, J-F. Aujol, and N. Papadakis, “Active contours without level sets,” pp. 2549–2552, 2012.
- [19] P. L. Combettes and J-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-point algorithms for inverse problems in science and engineering*, pp. 185–212. Springer, 2011.
- [20] J-J. Moreau, “Proximité et dualité dans un espace hilbertien,” *Bulletin de la Société mathématique de France*, vol. 93, pp. 273–299, 1965.